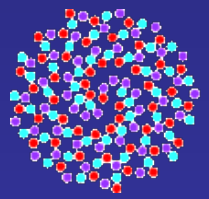
In Pursuit of a Working Model…

… The Curation Process

Catherine Lloyd



MAURICE WILKINS CENTRE
FOR MOLECULAR BIODISCOVERY

The Bioengineering Institute

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

Talk Outline

The problems with publishing

CellML as a solution

Creating and curating CellML models

Future goals

The Publishing Process…

model creation

x Error! x

translated for journal publication

x Error! x

the article is reviewed…

x Error! x

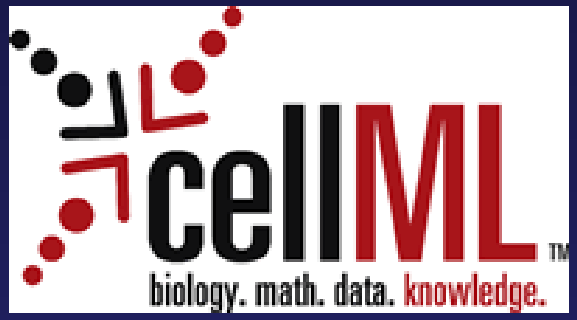
… and hopefully published

x Error! x

it’s read and interpreted

x Error! x

& finally implemented



CellML has evolved as a potential solution to the inconsistencies between computational and published models

It's an XML-based markup language which can be used to describe biological models

It's flexible structured to encourage model exchange, reuse and further development

Original Code ➡ Published Paper ➡ CellML

```
 %----------Calc the L-type Ca
current----------------------

            [CfCa,RevPCa]=
CalcConstantfield(Cai,Cao,2, Vm); %Ca
            [CfK,RevPK] =
CalcConstantfield(Ki,Ko,1, Vm);  %K
            [CfNa,RevPNa] =
CalcConstantfield(Nai,Nao,1, Vm);  %Na
            if (count ==1 && currenttime == 0)
               Va = -74.0078;
            else
               Va = Vm;
            end
            if (count ==0)
              [mcal, hcal,n] =
calcRateConst(1,Va,0,Cai,mcal,hcal,count,dt);
%Calc m and h
              ICaLNa =
(0.00005*PCAL*CfNa)*mcal*hcal;
              ICaLK = (0.001 * PCAL *
CfK)*mcal*hcal;
              %ICaLCa = (PCAL * CfCa*mcal*hcal)
%original
              ICaLCa = (PCAL * CfCa*mcal*hcal);
              ICaL = ICaLCa + ICaLK+ICaLNa;
            else
              ICaLNa =
(0.00005*PCAL*CfNa)*mcal*hcal;
              ICaLK = (0.001 * PCAL *
CfK)*mcal*hcal;
              ICaLCa = (PCAL * CfCa*mcal*hcal);
              ICaL = ICaLCa + ICaLK+ICaLNa;
              [mcal, hcal] =
calcRateConst(1,Va,0,Cai,mcal,hcal,count,dt);
%Calc m and h
            end
```

The cell membrane is modeled as a capacitor connected in parallel with variable resistances and batteries representing the different ionic currents and pumps. The electrophysiological behavior of a single cell can hence be described with the following differential equation (23)

$$\frac{dV}{dt} = -\frac{I_{ion} + I_{stim}}{C_m} \quad (1)$$

where V is voltage, t is time, I_{ion} is the sum of all transmembrane ionic currents, I_{stim} is the externally applied stimulus current, and C_m is cell capacitance per unit surface area.

Similarly, ignoring the discrete character of microscopic cardiac cell structure, a 2D sheet of cardiac cells can be modeled as a continuous system with the following partial differential equation (23)

$$\frac{\partial V}{\partial t} = -\frac{I_{ion} + I_{stim}}{C_m} + \frac{1}{\rho_x S_x C_m}\frac{\partial^2 V}{\partial x^2} + \frac{1}{\rho_y S_y C_m}\frac{\partial^2 V}{\partial y^2} \quad (2)$$

where ρ_x and ρ_y are the cellular resistivity in the x and y directions, S_x and S_y are the surface-to-volume ratio in the x and y directions, and I_{ion} is the sum of all transmembrane ionic currents given by the following equation

$$I_{ion} = I_{Na} + I_{K1} + I_{to} + I_{Kr} + I_{Ks} + I_{CaL} + I_{NaCa} + I_{NaK} + I_{pCa} + I_{pK} + I_{bCa} + I_{bNa} \quad (3)$$

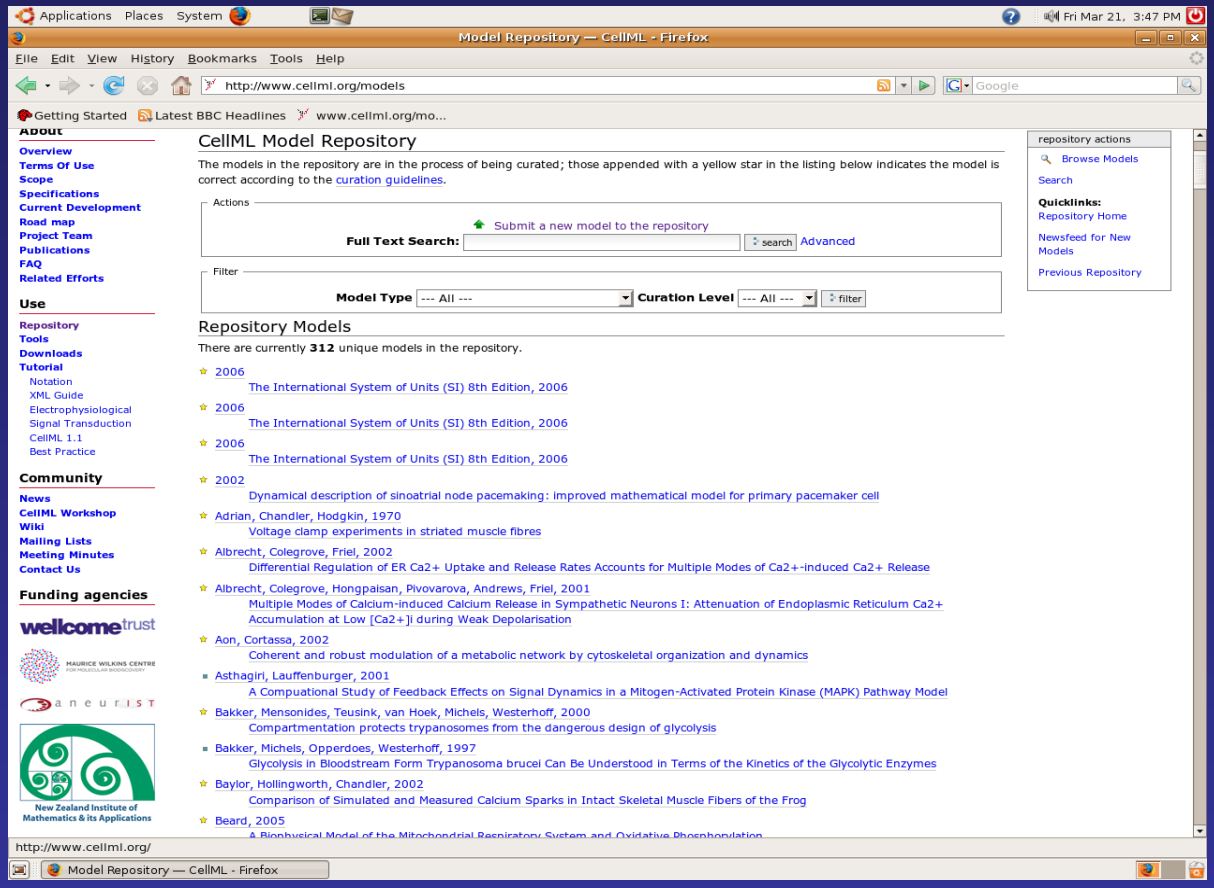
where I_{NaCa} is Na^+/Ca^{2+} exchanger current, I_{NaK} is Na^+/K^+ pump current, I_{pCa} and I_{pK} are plateau Ca^{2+} and K^+ currents, and I_{bCa} and I_{bK} are background Ca^{2+} and K^+ currents.

```
<component name="membrane">
    <variable units="millivolt" public_interface="out" cmeta:id="membrane_V"
name="V" initial_value="-86.2" />
    <variable units="joule_per_mole_kelvin" public_interface="out" name="R"
initial_value="8314.472" />
    <variable units="kelvin" public_interface="out" name="T" initial_value="310" />
    <variable units="coulomb_per_millimole" public_interface="out" name="F"
initial_value="96485.3415" />
    <variable units="microF" public_interface="out" name="Cm"
initial_value="0.185" />
    <variable units="micrometre3" public_interface="out" name="V_c"
initial_value="0.016404" />

    <variable units="millisecond" public_interface="in" name="time" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_K1" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_to" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_Kr" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_Ks" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_CaL" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_NaK" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_Na" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_b_Na" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_NaCa" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_b_Ca" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_p_K" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_p_Ca" />
    <variable units="picoA_per_picoF" public_interface="in" name="i_Stim" />

    <math xmlns="http://www.w3.org/1998/Math/MathML">
      <apply><eq />
        <apply><diff />
          <bvar><ci>time</ci></bvar>
          <ci>V</ci>
        </apply>
        <apply><times />
          <apply><divide />
            <apply><minus />
              <cn cellml:units="dimensionless">1</cn>
            </apply>
            <cn cellml:units="dimensionless">1</cn>
          </apply>
          <apply><plus /><ci>i_K1</ci>
            <ci>i_to</ci><ci>i_Kr</ci>
            <ci>i_Ks</ci><ci>i_CaL</ci>
            <ci>i_NaK</ci><ci>i_Na</ci>
            <ci>i_b_Na</ci><ci>i_NaCa</ci>
            <ci>i_b_Ca</ci><ci>i_p_K</ci>
            <ci>i_p_Ca</ci><ci>i_Stim</ci>
          </apply>
        </apply>
      </apply>
    </math>
  </component>
```

The Model Repository www.cellml.org/models

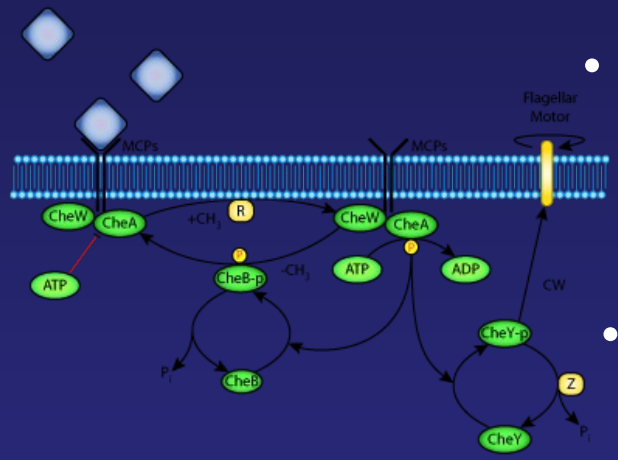
The CellML model repository began life as a set of examples to illustrate how the language could be applied and to test its features as it evolved

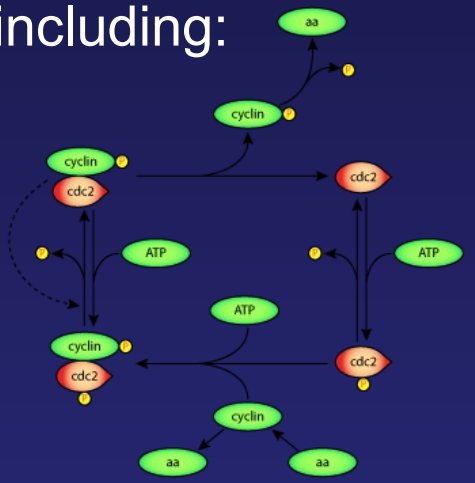
It later became a repository of previously published biological models, which were encoded in CellML based on the literature



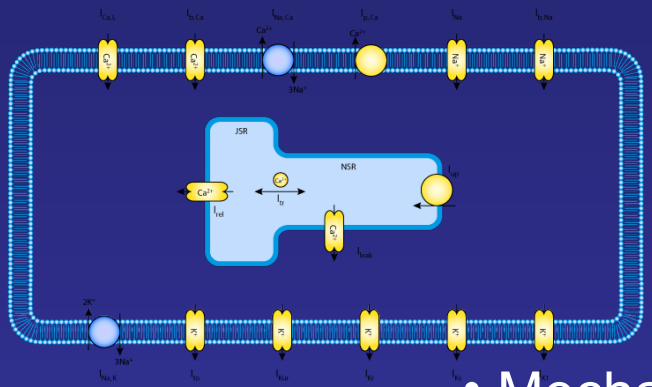
The Model Repository www.cellml.org/models

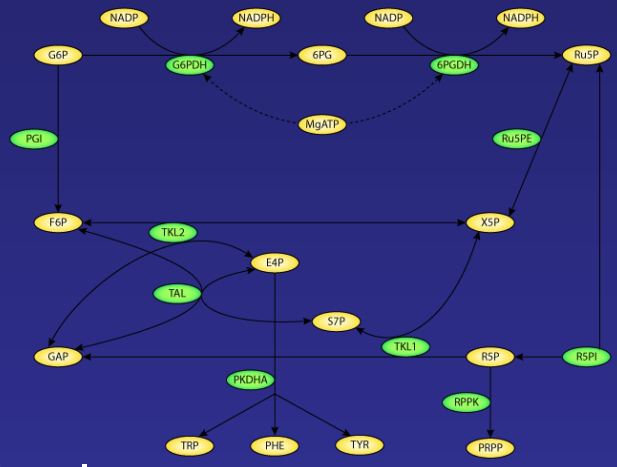
There are now >300 models in the repository including:





- Signal transduction pathways
- Metabolic pathways
- Electrophysiological models
- Immunological models
- Cell cycle models
- Muscle models
- Mechanical models & Constitutive laws





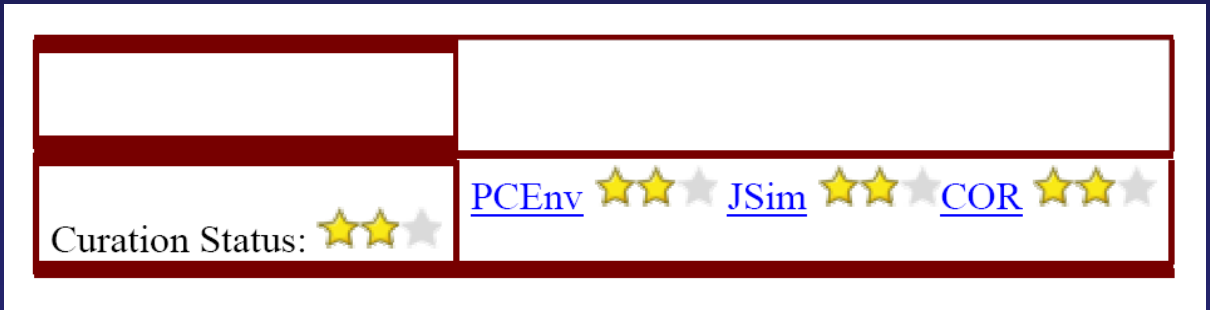
Model Creation & Curation

Initially, CellML models were written by hand in a text editor, and there was no way to test the models for consistency or completeness - they were simply checked by eye to see if they accurately represented the published model.

Since then tools, such as PCEnv and COR, have been developed to write and run CellML models, and these can also be used to curate the models in the repository.

Model Curation: The Theory

Of the 300 models in the repository ~½ have been curated

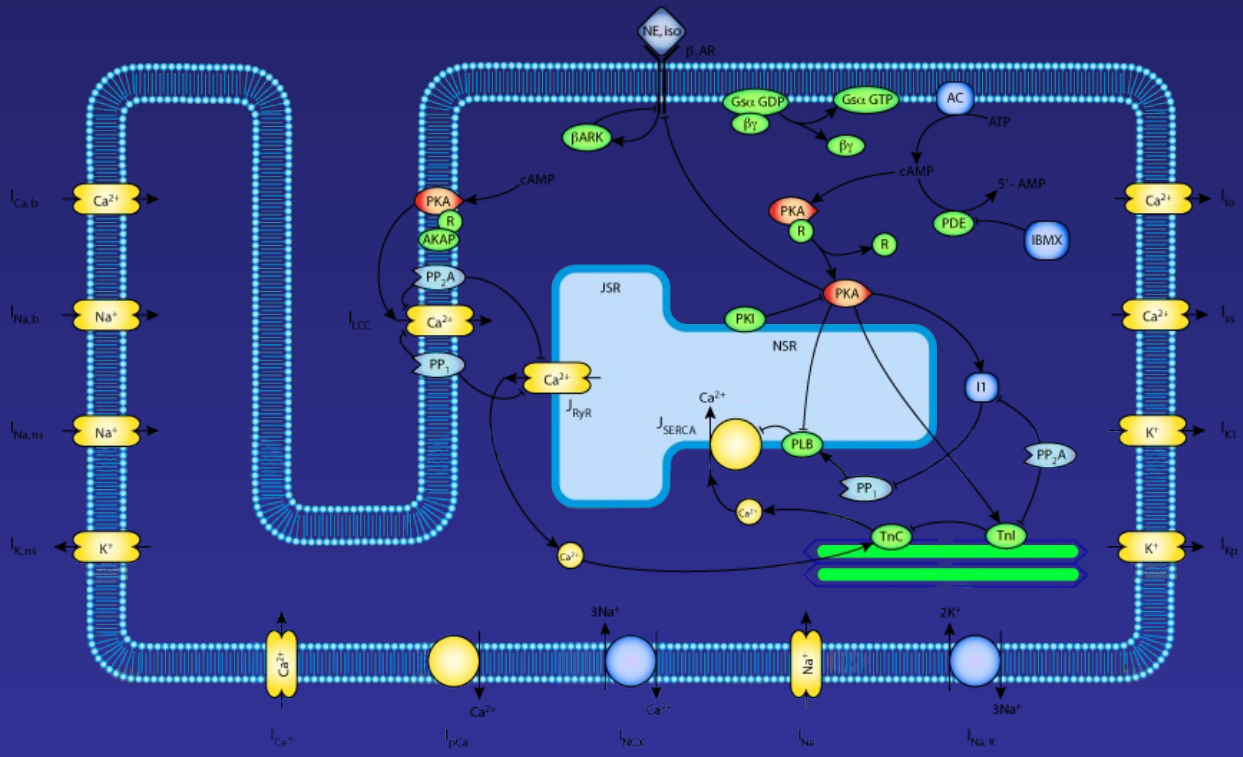


- Level 0: Not curated
- Level 1: Consistent with the published paper
- Level 2: Checked for (i) typos, (ii) unit consistency, (iii) completeness, (iv) not over-constrained, and (v) the model reproduces the published results;
- Level 3: the model is checked for the extent to which it satisfies physical constraints such as conservation of mass, momentum, charge, etc. (conducted by domain experts)

Point to Note

Model curation does not always imply “fixing” a model. A model may be valid CellML, and a true representation of the published paper, but there can be other reasons why the a CellML model won’t run

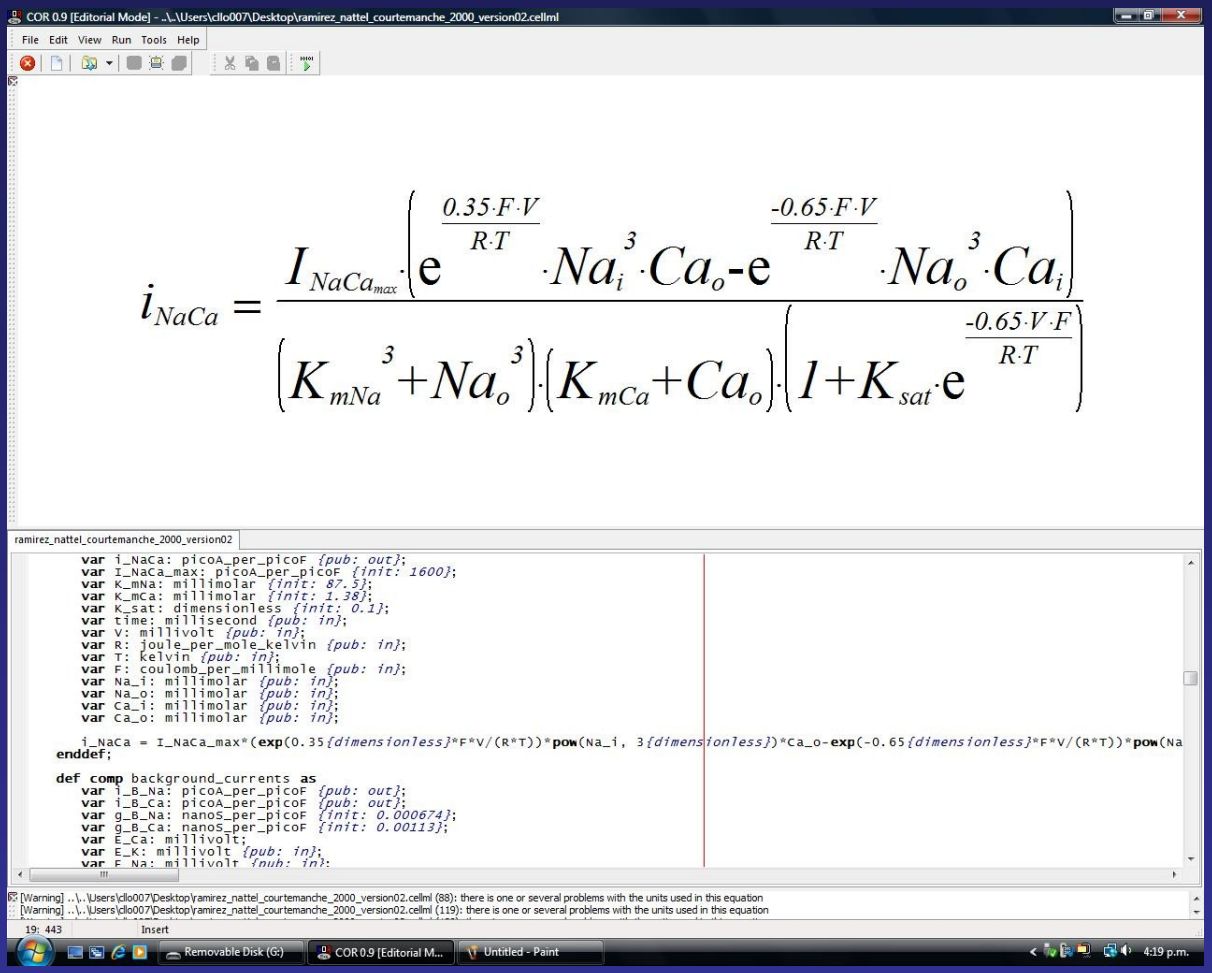
e.g. Saucerman *et al*. 2003 includes coupled, non-linear equations. The CellML model is valid, but currently the simulation tools PCEnv and COR are unable to handle this type of equation.



Model Curation – The Practice

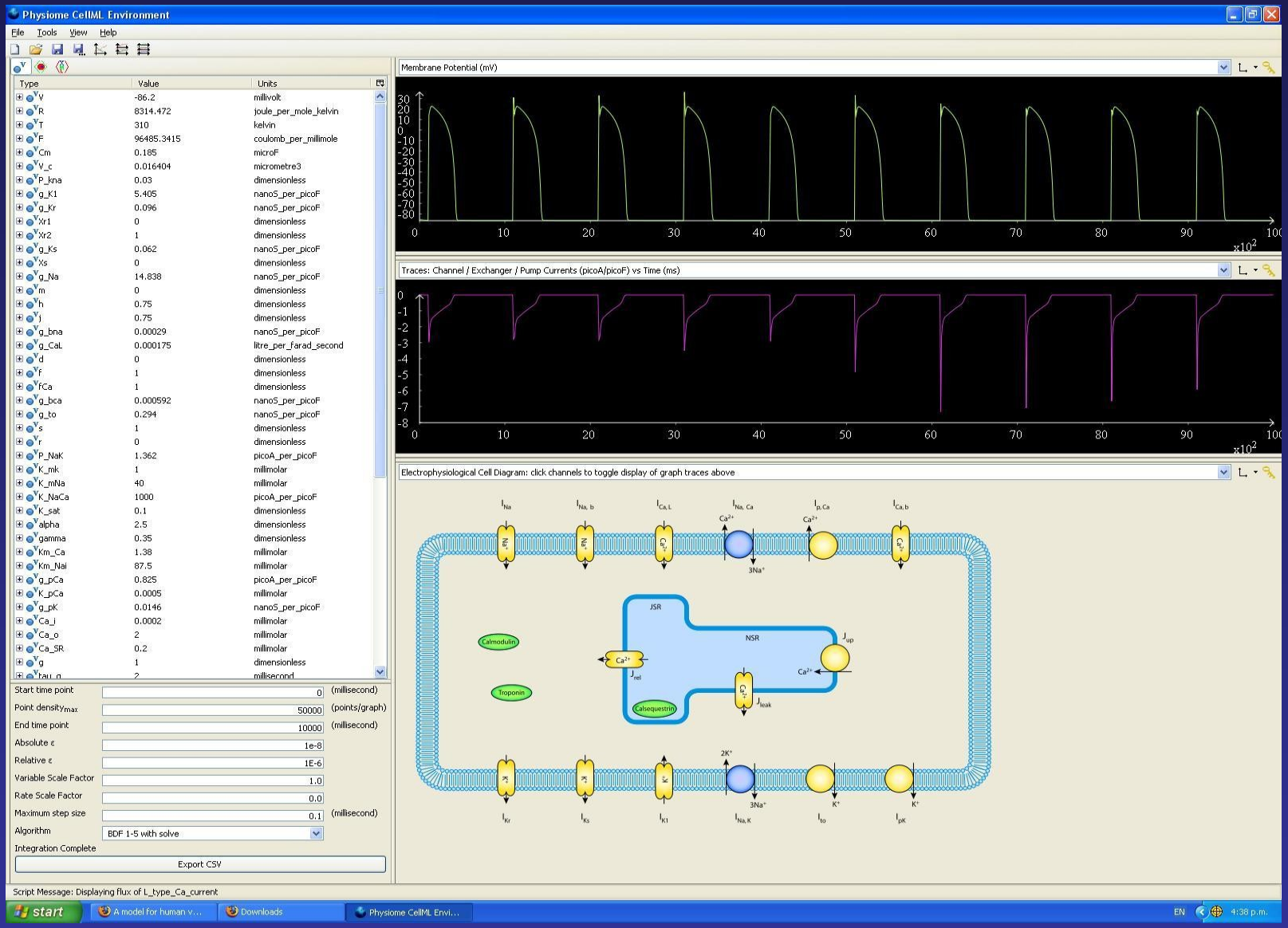
1) The CellML models are run in PCEnv and COR

& any obvious typos and unit inconsistencies are fixed



COR provides error messages & renders the MathML as readable equations

2) The simulation output is compared with the results in the published paper



3) Expert help is sought!

The model author is contacted and requests are made for missing parameters & equations, for general clarification and, where possible, the original code is retrieved

… often highlighting how error prone the publishing process is

The value of 'a' was indeed missing, it is 2.5218 (at temp = 286K). Also note that the value of delta-H for gamma should be 200240 rather than 200.24, this is an error in the Table (the period should have been a comma).

your guess is right: ko and k-o are 95 and 22/s, resp. They are not a function of voltage (as kv and k-v). **I'll fix the bug in the table and make it clearer for the print version.**

Thank you for your interest in our work and your careful reading of the paper. **Eq. 7 was printed wrong. Whilst proofreading the article for publication we found several misprints but we missed this one** (hopefully the only one). You are right, the last two iron terms in Eq. 7 should be in the ferrous form as in the pathway diagram (Fig. 1). Also, k8 and k8_ should have their units swopped over.

Communication with the model authors also frequently results in their positive feedback on our CellML efforts

That's sounds wonderful, I'm glad you were able to get the code to match the published results. I'm also glad to hear about the student interested in our 2007 model, **it's always nice to know that someone other than myself is interested in the models**.

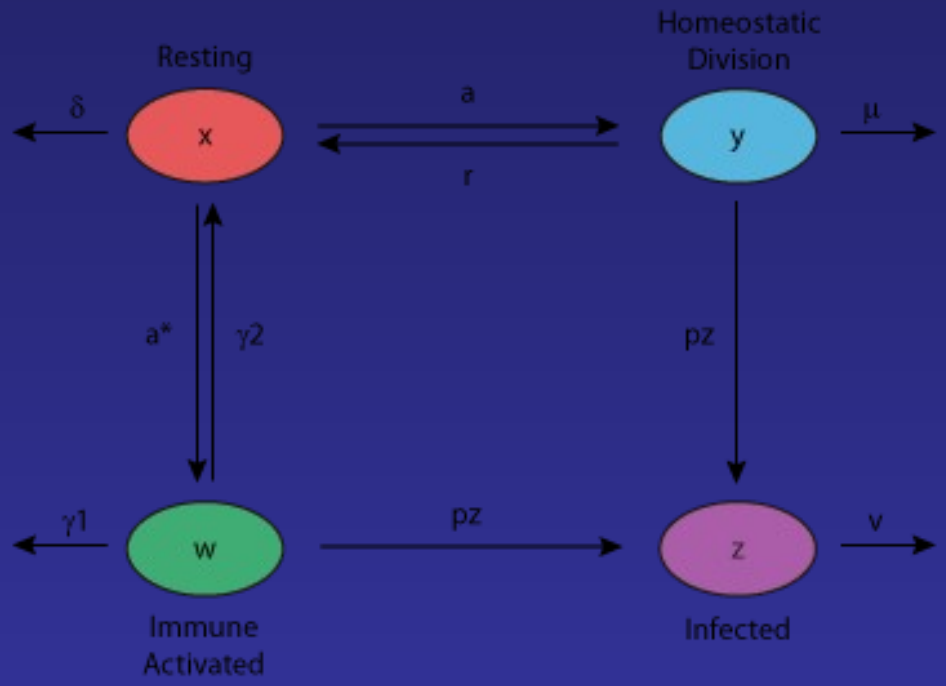
I was glad to see that you chose to simulate our model with your simulation set-up. I did not know about your system CellML. After my first look at it I thought it was an interesting tool, especially that it is coded in XML. **I will explore further the capabilities of your system more carefully...**

Parenthetically, I have been following the basic progress of some of the "ML" initiatives for the last few years, namely your effort and the SBML project. **I am delighted that you have chosen this model to use as an example in your repository. Quite exciting!**

I am very flattered that you've chosen to put our models on your web site (which I've just had a look at and it is great!).

And, just occasionally, we become the “experts”!

I'm currently teaching Yates et al 2007's PLos Med paper to a class of Zoology grad students and struggling mightily to replicate the authors' results. **The CellML model was extremely useful** in (1) identifying a typo in one of the equations in the published paper [the version on CellML agreed with common sense and the graphical model in the paper] and (2) identifying the scaling relation used to calculate one of the parameters (mu) from the values of the other parameters.



Future Goals

To complete the curation of all the models in the repository… ideally such that they recreate the results in the published paper

To encourage model developers (and journals) to publish their models in CellML code concurrent with their written paper

Acknowledgements

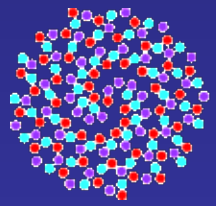
Curation Team: James Lawson & Penny Noble


Alan Garny & David Nickerson (for their enduring patience!)


Alan Garny, Andrew Miller & Justin Marsh (tool development)


Randall Britten & Poul Nielsen


All the model authors who have provided advice and feedback



MAURICE WILKINS CENTRE
FOR MOLECULAR BIODISCOVERY